PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| (51) International Patent Classification 6 : C12P 21/02, C12Q 1/68, C12N 15/11 | A1 | (11) International Publication Number: WO 99/43844 (43) International Publication Date: 2 September 1999 (02.09.99) |
|---|---|---|

(21) International Application Number: PCT/US99/04323

(22) International Filing Date: 26 February 1999 (26.02.99)

(30) Priority Data:

| | | |
|---|---|---|
| 09/032,684 | 27 February 1998 (27.02.98) | US |
| 09/185,115 | 3 November 1998 (03.11.98) | US |
| 09/197,889 | 23 November 1998 (23.11.98) | US |

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 09/197,889 (CIP)
Filed on 23 November 1998 (23.11.98)

(71) Applicant (for all designated States except US): THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; West 116th Street and Broadway, New York, NY 10027 (US).

(72) Inventor; and
(75) Inventor/Applicant (for US only): FISHER, Paul, B. [US/US]; 15 Gordon Place, Scarsdale, NY 10583 (US).

(74) Agent: WHITE, John, P.; Cooper & Dunham LLP, 1185 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AU, CA, JP, MX, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*

(54) Title: RECIPROCAL SUBTRACTION DIFFERENTIAL DISPLAY

(57) Abstract

This invention provides a method for identifying differentially expressed nucleic acids between two samples, comprising: a) selecting a first and second nucleic acid sample; b) producing libraries for the first and second nucleic acid sample; c) performing reciprocal subtraction between the libraries to produce two subtracted libraries; d) amplifying the two subtracted libraries; and e) comparing the two amplified subtracted libraries to identify differentially expressed nucleic acids. Also, this invention provides the above-described method, wherein the 3' primer used in the PCR amplification is an oligo dT 3' primer. This invention also provides the above-described methods, wherein the comparing of step e comprises using a gel to separate the nucleic acids from both of the libraries. This invention provides the isolated nucleic acid identified by the above-described methods, wherein the nucleic was not previously known to be differentially expressed between the two samples.

This invention also provides the above-described methods, wherein the comparing of step (e) comprises using a gel to separate the nucleic acids from both of the substracted libraries.

In an embodiment, the gel is a polyacrylamide gel. In another embodiment, the gel is an agarose gel.

This invention further provides the above-described methods, further comprising PCR amplifying the first and second nucleic acid samples.

This invention also provides the above-described methods, further comprising reamplifying differentially expressed bands.

This invention also provides the above-described methods, further comprising reamplifying differentially expressed nucleic acid.

In one method of reamplifying differentially expressed bands, differentially amplified bands from plasmids of each subtracted library were marked with an 18G needle through the film and cut out with a razor. The cut out differentially expressed bands can be reamplified (i.e. by PCR) and examined by reverse Northern and Northern blot analyses.

In addition, this invention provides the above-described methods, wherein the comparing of step (e) comprises comparing the band intensities of the two amplified differentially expressed nucleic acids.

In addition, this invention provides the above-described methods, wherein the nucleic acid samples are mRNA or cDNA derived from mRNA.

In addition, this invention provides the above-described methods, wherein the comparing of step (e) comprises comparing the quantities of the two amplified differentially expressed nucleic acids.

This invention further provides the above-described methods, wherein the differences in band intensity between the two subtracted libraries are electronically quantified.

This invention further provides the above-described methods, wherein the differences in the quantities of nucleic acid between the two subtracted libraries are electronically quantified.

In one embodiment, electronic quantification involves using a scanner to detect the bands. In a further embodiment, computer software, such as Corel Draw, can be used to determine the pixel intensity of the scanned image, thereby quantifying the band intensity.

Also, this invention provides the above-described methods, wherein the libraries of step (b) are constructed with λ-ZAP cDNA library kits. One skilled in the art would recognize that any cDNA library would be suitable.

This invention provides the isolated nucleic acid identified by the the above-described methods, wherein the nucleic acid was not previously known.

This invention also provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 12 (AI 144569).

In addition, this invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid

is the nucleic acid designated PSGen 13 (Accession No. AI 144570).

This invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 23.

This invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 24.

This invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 25.

This invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 26 (Accession No. AI 144571).

This invention also provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 27 (Accession No. AI 144572).

This invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 28 (AI 144573).

This invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PSGen 29 (AI 144574).

This invention provides the above-described isolated nucleic acid, wherein the isolated nucleic acid is the nucleic acid designated PEGen 13 (AI 144564).

## Second Series of Experiments

Presently described is a RSDD approach that efficiently and consistently reduces the complexity of DDRT-PCR and results in the identification and cloning of genes displaying anticipated differential expression. The model used for RSDD was an adenovirus-transformed rat embryo cell line, E11, that acquires an aggressive oncogenic progression phenotype when injected into athymic nude mice and reestablished in cell culture (E11-NMT)(6,26,27). Injection of E11 cells into nude mice results in tumors in 100% of animals with a tumor latency time of approximately 35 to 40 days, whereas E11-NMT cells form tumors in 100% of nude mice with a tumor latency time of 15 to 20 days (6,26,27). Additionally, E11 cells form colonies in agar with an efficiency of ~3 %, whereas E11-NMT display an agar cloning efficiency of >30% (6,26,27). The increased tumorigenicity and enhanced anchorage independence phenotypes are key indicators of tumor progression in the E11/E11-NMT model system (6,26,27). RSDD has resulted in the identification and cloning of genes displaying elevated expression in progressed tumor cells (progression elevated gene, PEGen) and suppressed expression in progressed tumor cells (progression suppressed gene, PSGen).

## MATERIALS AND METHODS

**RNA isolation and cDNA library construction.** Total RNA from E11 and E11-NMT cells was isolated by the guanidinium isothiocyanate/CsCl centrifugation procedure and poly(A)$^+$ RNA was purified with oligo(dT) cellulose chromatography(5). Two λ-ZAP cDNA libraries from E11 and E11-NMT mRNAs were constructed with λ-ZAP cDNA library kits (Stratagene) following the manufacturer's protocol. Reciprocal subtraction between E11 and E11-NMT libraries was performed and two subtracted cDNA libraries (E11 minus E11-NMT and E11-NMT minus E11) were constructed as

described(5,6). Plasmid cDNA libraries from the subtracted λ-ZAP cDNA libraries were obtained by in vivo excision following the manufacturer's protocol (Stratagene) and the plasmids were isolated with Qiagen columns (Qiagen, Chatsworth, CA.).

**RSDD methodology.** The purified plasmids of reciprocally subtracted cDNA libraries were directly subjected to differential display as in Liang et al. (28) with minor modifications. The plasmids of reciprocally subtracted cDNA libraries were PCR-amplified with the combination of three single-anchor 3' primers ($T_{13}A$, $T_{13}C$ or $T_{13}G$) and 18 arbitrary 5' 10-mer primers obtained from Operon Technology Inc. (Alameda, CA. OPA 1-20 except OPA1 and 3). The 20 $\mu$l PCR reaction consisted of 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 2 $\mu$M each dNTP, 0.2 $\mu$M 5' arbitrary primer, 1 $\mu$M 3' anchor primer, 50 ng of plasmid of a subtracted library, 10 $\mu$Ci $\alpha$-$^{35}$S-dATP (3,000 Ci/mmol from Amersham) and 1 unit of Taq DNA polymerase (Gibco/BRL). The parameters of PCR were 30 sec at 95°C, 40 cycles of 30 sec at 95°C, 2 min at 40°C and 30 sec at 72°C and additional 5 min. at 72°C. After the cycling, 10 $\mu$l of 95% formamide, 0.05% bromophenol blue and 0.05% xylene cyanol were added to each PCR reaction. The mixture was heated at 95°C for 2 min and separated in a 5% denaturing DNA sequencing gel maintained at 50°C. PCR reactions of plasmids from each subtracted library in a primer set were run side by side. Differentially amplified bands from plasmids of each subtracted library were marked with 18G needle through the film and cut out with a razor. The gel slice was put in 100 $\mu$l TE (pH 8.0) and incubated at 4°C overnight. After the incubation, the mixture was boiled for 5 min and microcentrifuged for two min. The supernatant was collected and stored at -20°C until reamplification. The band extract was reamplified with the same cycling parameters in a 50 $\mu$l reaction consisting of 10 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM

$MgCl_2$, 20 $\mu$M each dNTP, 0.2 $\mu$M 5' arbitrary primer, 1 $\mu$M 3' anchor primer, 5$\mu$l of band extract and 2.5 units of Taq DNA polymerase (Gibco/BRL).

**Reverse Northern Blotting Procedure.** Differential expression of the reamplified DNA fragment was scrutinized by reverse Northern and Northern blot analyses. In reverse Northern analysis, after confirmation in a 1% agarose gel, the reamplified DNA fragment (10 $\mu$l of PCR reaction) was mixed with 90 $\mu$l TE and spotted on a positively charged Nylon membrane (Boehringer Mannheim) with a 96-well vacuum manifold. The membrane was soaked with denaturing and neutralizing solution successively, and the spotted DNA was crosslinked to the membrane with a UV crosslinker (Stratagene). $^{32}$P-labeled first strand cDNA was prepared by reverse transcription of total RNA. After heating at 70°C for 10 min and quenching on ice for two min, 0.4 $\mu$M each $T_{13}A$, $T_{13}G$ and $T_{13}C$ and 10 $\mu$g total RNA mixture was added with 50 mM Tris-HCl, (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM dATP, 0.5 mM dGTP, 0.5 mM dTTP, 0.02 mM dCTP, 0.5 $\mu$l RNase inhibitor (Gibco/BRL), 100 $\mu$Ci dCTP (3,000 Ci/mmol from Amersham) and 200 units Superscript RT II (Gibco/BRL) in a final 25 $\mu$l reaction. The reaction mixture was incubated at 42°C for one hour and at 37°C for 30 min after addition of 2 $\mu$l of RNase H (10 units, Gibco/BRL). The membrane was hybridized at 42°C overnight in a 50% formamide hybridization solution. The hybridized membrane was washed at room temperature for 15 min with 2X standard saline citrate containing 0.1% SDS twice and at 55°C for at least one hour with 0.1X Standard Saline Citrate containing 0.1% SDS, successively. The membrane was probed with the $^{32}$P-labeled cDNA of E11, striped off and probed with $^{32}$P-labeled cDNA of E11-NMT. The signal intensity of each spot was normalized against that of glyceraldehyde-3-phosphate dehydrogenase and compared between E11 and E11-NMT. Reamplified DNA

# FIG. 35B

**PSGen 13 cDNA Sequence**

```
GGCACGAGCT CTCCTCGTCC CCTCCCTTCT CCACTGCAGC CTTTCTCTTA
GCCCGAACCA CTTCCTTCTT CTGCTTGTTC CTCCCTAGGG CGCGGAAGCT
GAGTGCAGGG TTCAGACCCA CGCGGCGAGC AGCTCTTCAG TGAAGAAGGA
AGCAATCGGA GGGTCAGCAA TGAACGTGGA GCATGAGGTT AACCTCCTGG
TGGAGGAAAT TCATCGTCTG GGTTCCAAAA ATGCCGATGG GAAACTGAGT
GTGAAGTTTG GGGTCCTCTT CCAAGACGAC AGATGTGCCA ATCTCTTTGA
AACCGTTGGT GGGAACTCTG AAAGCCCGCA AAACGAAGGA AGATTGTTAC
GTACGCAGAA GAGCTGCTTT TGCAAGGTGT TCATGATGAT GTTGACATTG
TATTGCTGCA AGATTAATGT GGTTTGCAGA TCTGGGGGTA TCTGGTAAAC
TGGAATAATT AAGTTAAAGG ACAAACATGA AGTTCCTTAT GTATTTTTAT
AGACCTTTGT AAACAAAAGG GGACTTGTTG AGAAGTCCTG TTTTTATACC
TTGGAGCAAA ACATTACAAT GTAAAAATAA ACAAAACCTG TTATTTTTTT
TTTCTTAAGA AGGTAATCGG GAGACGTAGG CAATAAAATG TTTTCAGAGG
TGCGAAAAAG CTTTTGTTTT CTTAAACCAT TCTTAGTCTC TGCCACACTT
GACACTCCGT CAAAGTGAGA AGCGAACTAA AGACCAACTG CGGTGGAAAA
TATTATGTTT ATGTAATAAA AAAAAATCAT GTAAAAAAAA AAAAAAAAAA
```

**PSGen 13 Protein Sequence**

```
MNVEHEVNLL VEEIHRLGSK NADGKLSVKF GVLFQDDRCA NLFETVGGNS
ESPQNEGRLL RTQKSCFCKV FMMMLTLYCC KINVVCRSGG IW·
```

WEST

15. The method of claim 1 or 2, further comprising reamplifying differentially expressed nucleic acids.

16. The method of claim 1 or 2, wherein the comparing of step (e) comprises comparing the quantities of the two amplified differentially expressed nucleic acids.

17. The method of claim 1 or 2, wherein differences in the quantities of nucleic acid between the two subtracted libraries are electronically quantified.

18. The method of claim 1 or 2, wherein the libraries of step (b) are constructed with λ-ZAP cDNA library kits.

19. The isolated nucleic acid identified by the method of claim 1 or 2, wherein the nucleic acid was not previously known.

20. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 12.

21. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 13.

22. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 23.

23. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 24.

24. The isolated nucleic acid of claim 19, wherein the

isolated nucleic acid is the nucleic acid designated PSGen 25.

25. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 26.

26. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 27.

27. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 28.

28. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PSGen 29.

29. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 13.

30. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 14.

31. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 15.

32. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 24.

33. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated

PEGen 28.

34. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 32.

35. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 42.

36. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 43.

37. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 44.

38. The isolated nucleic acid of claim 19, wherein the isolated nucleic acid is the nucleic acid designated PEGen 48.

39. The isolated nucleic acid molecule of claim 19 which comprises:

(a) one of the nucleic acid sequences as set forth in Figure 35;

(b) a sequence being degenerated to a sequence of (a) as a result of the genetic code;

(c) a sequence encoding one of the amino acid sequences as set forth in Figure 35.

(d) a sequence of at least 12 nucleotides capable of specifically hybridizing to the sequence of (a), (b) or (c)